Amendments to the Specification

In the Specification:

Please cancel the existing sequence listing for the above-identified application, and replace it with the substitute sheets appended hereto. Please renumber the remaining pages accordingly. A computer readable copy of the substitute sequence listing is forwarded herewith.

At page 1, after the title and prior to the "Background of the Invention" section, please insert the following:

--Cross-reference to Related Application

This application is a divisional of U.S. Application No. 08/841,636, filed April 30, 1997 (pending), which is a continuation of International Application No. PCT/FI96/00550, filed October 17, 1996, and a continuation-in-part of U.S. Application No. 08/732,181, filed October 16, 1996 (abandoned), which claims the benefit of U.S. Provisional Application Nos. 60/005,335, filed October 17, 1995; 60/007,926, filed December 4, 1995; and 60/020,840, filed June 28, 1996.

Please amend the following paragraphs/sections as follows.

Please amend the paragraph beginning on page 7, line 21, as follows:

Figure 17 shows amino acid sequence data derived from sequencing the 20K-cellulase described in the exemplary material herein. Sequence 429 (SEQ ID NO:1) is from the N terminus of the protein and the other sequences are from internal tryptic peptides. Sequence #430 corresponds to SEQ ID NO: 2; sequence #431 corresponds to SEQ ID NO: 3; sequence #432 corresponds to SEQ ID NO: 4; sequence #433 corresponds to SEQ ID NO: 5; sequence #439 corresponds to SEQ ID NO: 6; fr 9 corresponds to SEQ ID NO: 7; fr 14 corresponds to SEQ ID NO: 8; fr 16 corresponds to SEQ ID NO: 9; fr 17 corresponds to SEQ ID NO: 10; fr 28 corresponds to SEQ ID NO: 11 and fr 30 corresponds to SEQ ID NO: 12.

Please amend the paragraph beginning on page 7, line 28, as follows:

Figure 19 (A and B) shows the DNA sequence of the 20K-cellulase gene (SEQ ID NO: 30). The arrow indicates the predicted signal peptidase processing site.

Please amend the paragraph beginning on page 8, line 4, as follows:

Figure 21 ~~(A and B)~~ (A, B and C) shows the DNA sequence of the 50K-cellulase gene (SEQ ID NO: 32). The arrow indicates the predicted signal peptidase processing site.

Please amend the paragraph beginning on page 8, line 9, as follows:

Figure 23 ~~(A and B)~~ (A, B and C) shows the DNA sequence of the 50K-cellulase B gene (SEQ ID NO: 34). The arrow indicates the predicted signal peptidase processing site.

Please amend the paragraph beginning on page 8, line 16, as follows:

Figure 27 shows the DNA sequence of the protein-with-CBD cellulase gene (SEQ ID NO: 36) in pALK1230.

Please amend the paragraph beginning on page 18, line 3, as follows:

A nucleic acid molecule encoding a polypeptide having the enzymatic activity of a cellulase, selected from the group consisting of:

(a) nucleic acid molecules encoding a polypeptide comprising the amino acid sequence as depicted in Figure 19 (SEQ ID NO: 31) or 21 (SEQ ID NO: 33);

(b) nucleic acid molecules encoding a polypeptide comprising the amino acid sequence as depicted in Figure 23 (SEQ ID NO: 35) or 27 (SEQ ID NO: 37);

(c) nucleic acid molecules comprising the coding sequence of the nucleotide sequence as depicted in Figure 19 (SEQ ID NO: 30) or 21 (SEQ ID NO: 32);

(d) nucleic acid molecules comprising the coding sequence of the nucleotide sequence as depicted in Figure 23(SEQ ID NO: 34) or 27 (SEQ ID NO: 36);

(e) nucleic acid molecules encoding a polypeptide comprising the amino acid sequence encoded by the DNA insert contained in DSM 11024, DSM 11012, DSM 11025 or DSM 11014;

(f) nucleic acid molecules encoding a polypeptide comprising the amino acid sequence encoded by the DNA insert contained in DSM 11026, DSM 11011, DSM 11013 or DSM 11027;

(g) nucleic acid molecules comprising the coding sequence of the DNA insert contained in DSM 11024, DSM 11012, DSM 11025 or DSM 11014;

(h) nucleic acid molecules comprising the coding sequence of the DNA insert contained in DSM 11026, DSM 11011, DSM 11013 or DSM 11027;

(i) nucleic acid molecules hybridizing to a molecule of any one of (a), (c), (e) or (g); and

(j) nucleic acid molecules the coding sequence of which differs from the coding sequence of a nucleic acid molecule of any one of (a) to (i) due to the degeneracy of the genetic code; and

(k) nucleic acid molecules encoding a polypeptide having cellulase activity and having an amino acid sequence which shows at least 80% identity to a sequence as depicted in Figure 19 (SEQ ID NO: 31), 21 (SEQ ID NO: 33), 23 (SEQ ID NO: 35) or 27 (SEQ ID NO: 37).

Please amend the paragraph beginning on page 48, line 16, as follows:

Amino acid sequences of tryptic peptides derived from 20K-cellulases are shown in Figure 17. Sequence #429 corresponds to SEQ ID NO: 1; sequence #430 corresponds to SEQ ID NO: 2; sequence #431 corresponds to SEQ ID NO: 3; sequence #432 corresponds to SEQ ID NO: 4; sequence #433 corresponds to SEQ ID NO: 5; sequence #439 corresponds to SEQ ID NO: 6; fr 9 corresponds to SEQ ID NO: 7; fr 14 corresponds to SEQ ID NO: 8; fr 16 corresponds to SEQ ID NO: 9; fr 17 corresponds to SEQ ID NO: 10; fr 28 corresponds to SEQ ID NO: 11 and fr 30 corresponds to SEQ ID NO: 12.

Please amend the paragraph beginning on page 53, line 14, as follows:

Table IX

Sequences of peptides isolated from the 50K-cellulase (uncertain residues in lower case)

#507 (SEQ ID NO: 13)	VYLLDETEHR
#509 (SEQ ID NO: 14)	XXLNPGGAYYGT
#563 (SEQ ID NO: 15)	MsEGAECEYDGVCDKDG
#565 (SEQ ID NO: 16)	NPYRVXITDYYGNS
#603 (SEQ ID NO: 17)	DPTGARSELNPGGAYYGTGYXDAQ
#605 (SEQ ID NO: 18)	XXVPDYhQHGVda
#610 (SEQ ID NO: 19)	NEMDIXEANSRA
#611 (SEQ ID NO: 20)	LPXGMNSALYLSEMDPTGARSELNP
#612 (SEQ ID NO: 21)	VEPSPEVTYSNLRXGEIXgXF
#619 (SEQ ID NO: 22)	DGCGWNPYRVvITtDYYnN
#620 (SEQ ID NO: 23)	LPCGMXSALY
#621 (SEQ ID NO: 24)	ADGCQPRTNYIVLDdLlHPXXQ

Please amend the paragraph beginning on page 55, line 4, as follows:

Table X

Sequences of peptides isolated from the 50K-cellulase B (uncertain residues in lower case)

#534 (SEQ ID NO: 25)	vGNPDFYGK
#535 (SEQ ID NO: 26)	FGPIGSTY
#631 (SEQ ID NO: 27)	LSQYFIQDGeRK
#632 (SEQ ID NO: 28)	FTVVSRFEENK
#636 (SEQ ID NO: 29)	HEYGTNVGSRFLYLMNGPDK

Please amend the paragraph beginning on page 67, line 4, as follows:

To amplify the 20K-cellulase gene by polymerase chain reaction (PCR), a pair of degenerate primers based on the peptide sequences (Figure 17)(SEQ ID NOS: 1-12) was synthesized. Primer 1 (429-32)(SEQ ID NO: 38) was derived from the amino acids #8-14 of the N-terminal peptide #429 (Figure 17)(SEQ ID NO: 1), and primer 2 (fr28-16)(SEQ ID NO: 39) was designed as the antisense strand for the amino acids #2-8 of the peptide fr28 (Figure 17)(SEQ ID NO: 11). Additional *Eco*R1 restriction sites were added at the 5'-termini to facilitate the cloning of the amplified fragment.

Please amend the paragraph beginning on page 67, line 12, as follows:

Primer 1 (429-32)(SEQ ID NO: 38)

Please amend the paragraph beginning on page 67, line 17, as follows:

Primer 2 (fr28-16)(SEQ ID NO: 39)

Please amend the paragraph beginning on page 69, line 7, as follows:

The insert (594 bp) in pALK549 was found to encode the majority of the 20K-cellulase derived peptide (Figure 17)(SEQ ID NOS: 1-12). The PCR amplified DNA (in addition to the primers) corresponds to the nucleotides 175-716 in Figure 19(A and B)(SEQ ID NO: 30).

Please amend the paragraph beginning on page 70, line 18, as follows:

The *Melanocarpus albomyces* DNA in pALK1221 was sequenced as described in Example 19. The DNA sequence encoding the *Melanocarpus albomyces* 20K-cellulase is

shown in Figure 19 (A and B)(SEQ ID NO: 30). The sequence is 936 bp in length, and has an open reading frame (ORF) coding for 235 amino acids; the gene has two introns. The putative signal peptide processing site is after alanine-21, and the N-terminus of the mature protein begins at alanine-22, as suggested by the peptide sequencing results (Figure 17, peptide #429)(SEQ ID NO: 1). The ORF predicts a protein with a molecular weight of 25.0 kDa for the full-length preprotein, and 22.9 kDa for the mature protein. This is in good agreement with the results obtained from the protein purification work (Example 10). These results also verify that the about 35 kDa protein detected previously with the 20K-cellulase antiserum (Example 10) is a different gene product than the 20K-cellulase.

Please amend the paragraph beginning on page 71, line 14, as follows:

The peptides derived from the 50K-cellulase (Table IX) shared some homology towards *Humicola grisea* endonuclease I (DDBJ:D63516). To amplify the 50 K-cellulase gene by polymerase chain reaction (PCR) a pair of degenerate primers based on the peptide sequences (Table IX)(SEQ ID NOS: 13-24) was synthesized. Primer 1 (507-128)(SEQ ID NO: 40) was derived from the amino acids #5-10 of the peptide #507 (Table IX)(SEQ ID NO: 13), and primer 2 (509-rev)(SEQ ID NO: 41) was designed as the antisense strand for the amino acids #4-9 of the peptide 509 (Table IX)(SEQ ID NO: 14). The order of the two peptides in the protein-and the corresponding sense-antisense nature of the primers-was deduced from comparison with the *Humicola grisea* endonuclease I.

Please amend the paragraph beginning on page 71, line 23, as follows:

Primer 1 (507-128)(SEQ ID NO: 40)

Please amend the paragraph beginning on page 72, line 1, as follows:

Primer 2 (509-rev)(SEQ ID NO: 41)

Please amend the paragraph beginning on page 73, line 10, as follows:

The insert (161 bp) in pALK1064 was sequenced as described in Example 19, and was found to contain an ORF, which predicted a peptide homologous to *Humicola grisea* endoglucanase I (DDBJ:D63516). The ORF also encoded the peptide #612 (Table IX)(SEQ ID NO: 21) from the purified 50K-cellulase. The PCR amplified DNA (in addition to the primers) corresponds to the nucleotides 404-530 in Figure 21(SEQ ID NO: 32).

Please amend the paragraph beginning on page 74, line 15, as follows:

The DNA encoding the *Melanocarpus albomyces* 50K-cellulase is shown in Figure 21 ~~(A and B)~~(A, B and C)(SEQ ID NO: 32). The sequence reveals an ORF of about 1363 bp in length, interrupted by one intron. The ORF codes for 428 amino acids. The predicted protein has a molecular weight of 46.8 kDa and after signal peptide cleavage of 44.8 kDa. All the peptides in Table IX (SEQ ID NOS: 13-24) are found in the predicted protein sequence ~~(Figure 2)~~(Figure 21)(SEQ ID NO: 33.), although some amino acids identified with uncertainty during the peptide sequencing proved to be incorrect. The protein shows homology to *Humicola grisea* endoglucanase I (DDBJ:D63516).

Please amend the paragraph beginning on page 74, line 26, as follows:

The peptides derived from the 50K-cellulase B (Table X)(SEQ ID NOS: 25-29) shared some homology towards *Humicola grisea* cellobiohydrolase I (DDBJ:D63515). To amplify the 50K-cellulase B gene by polymerase chain reaction (PCR) a pair of degenerate

primers based on the peptide sequences (Table X)(SEQ ID NOS: 25-29) was synthesized. Primer 1 (636)(SEQ ID NO: 42) was derived from the amino acids #1-5 of the peptide #636 (Table X)(SEQ ID NO: 29) (the first amino acids was guessed to be lysine, because the peptide was isolated after digestion with a protease cleaving after lysines), and primer 2 (534-rev)(SEQ ID NO: 43) was desinged as the antisense strand for the amino acids #3-8 of the peptide #534 (Table X)(SEQ ID NO: 25). The order of the two peptides in the protein-and the corresponding sense-antisense nature of the primers-was deduced from comparison with the *Humicola grisea* cellobiohydrolase I.

Please amend the paragraph beginning on page 75, line 8, as follows:

Primer 1 (636)(SEQ ID NO: 42)

Please amend the paragraph beginning on page 75, line 11, as follows:

Primer 2 (534-rev)(SEQ ID NO: 43)

Please amend the paragraph beginning on page 76, line 21, as follows:

The insert in pALK1224 was sequenced as described in Example 19, and was found to contain an ORF encoding the whole peptide #636 (SEQ ID NO: 29) from 50K-cellulase B (Table X). The ORF predicted a peptide homologous to *Humicola grisea* cellobiohydrolase I (DDBJ:D63515). The PCR amplified DNA (in addition to the primers) corresponds to the nucleotides 371-1023 in Figure 23 (A, B and C)(SEQ ID NO: 34).

Please amend the paragraph beginning on page 77, line 29, as follows:

Part of the inserts in pALK1229 and pALK1236 were sequenced as described in Example 19. The DNA encoding the *Melanocarpus albomyces* 50K-cellulase B is shown in ~~Figure 23 (A and B)~~(Figure 23A, B and C)(SEQ ID NO: 34). The sequence reveals an ORF of 1734 bp in length interrupted by five introns. The ORF codes for 452 amino acids. The predicted protein has a molecular weight of 49.9 kDa and after signal peptide cleavage of 47.6 kDa. All the peptides in Table X (SEQ ID NOS: 25-29) are found in the predicted protein sequence ~~(Figure 23A and B)~~(Figure 23A, B and C)(SEQ ID NO: 35)~~,~~ although some amino acids identified with uncertainty during the peptide sequencing proved to be incorrect. The predicted protein shows homology to *Humicola grisea* cellobiohydrolase I (DDBJ:D63515) and other cellobiohydrolases. However, 50K-cellulase B has the surprising feature that it does not harbor the cellulose binding domain (CBD) and its linker, which is characteristic to *Humicola grisea* cellobiohydrolase I and many other cellobiohydrolases.

Please amend the paragraph beginning on page 79, line 18, as follows:

Part of the insert in pALK1230 was sequenced as described in Example 19. The DNA appears not to encode the 20K-cellulase, but codes for a protein homologous to several cellulases, particularly at the cellulose binding domain (CBD) area. Thus the gene product very likely has high affinity towards cellulosic material, and therefor this gene product was designated as protein-with-CBC. The sequence is shown in Figure 27 (SEQ ID NO: 36).

Please amend the paragraph beginning on page 79, line 24, as follows:

PCR reactions with the primers 636 (SEQ ID NO: 42) and 534-rev (SEQ ID NO: 43) (Example 23) were performed with the DNA from the 19 lambda clones as templates. One lambda clone, lambda-3, gave a band about 700 bp in size, similar to that in Example 23 when ALKO4237 chromosomal DNA was used as a template. This clone had originally been picked by the *Trichoderma cbh1* probe. The lambda DNA was digested with several restriction endonucleases, and hybridized with the 50K-cellulase B specific probe. The clone showed similar restriction enzyme pattern as the 3 clones in Example 24. It is concluded that lambda -3 also carries the 50K-cellulase B gene.

Please amend the paragraph beginning on page 85, line 3, as follows:

**T. reesei cbh1* (cellobiohydrolase 1) promoter: The promoter is from *Trichoderma reesei* VTT-D-80133 (Teeri et al. (1983) The molecular cloning of the major cellulase gene from *Trichoderma reesei*. *Bio/Technology* 1: 696.). The 2.2 kb *Eco*R1-*Sac*II fragment (Karhunen et al. (1993) High frequency one-step gene replacement in *Trichoderma reesei*. I. Endoglucanase I overproduction. *Mol. Gen. Genet.* 241:515) was used in the construct. The sequence of the promoter area preceding the ATG was published by Shoemaker et al. (1983) Molecular cloning of exo-cellobiohydrolase from *Trichoderma reesei* strain L27. *Bio/Technology* 1.691.). The last 15 nucleotides of the *T. reesei* L27 *cbh1* promoter (the *Sac*II site is underlined) are CCGCGGACTGGCATC (SEQ ID NO: 44) (Shoemaker et al. 1983). The *cbh*1 promoter from the *T. reesei* strain VTT-D-80133 has been sequenced at Alko Research Laboratories, and one nucleotide difference in the DNA sequence has been noticed within the above mentioned region. In the *T. reesei* strain VTT-D-80133 the

sequence preceding the ATG is CCGCGGACTG/C/GCATC (SEQ ID NO: 45) (the *Sac*II site is underlined, the additional cytosine in the DNA sequence is between the slashes).

Please amend the paragraph beginning on page 85, line 28, as follows:

**Melanocarpus albomyces* 20K-cellulase gene: The nucleotide sequence and deduced amino acid sequence of the 20K-cellulase gene encoding a 20 kDa cellulase is presented in Example 20 (Figure 19)(SEQ ID NOS: 30-31). A 0.9 kb fragment beginning from ATG-codon was used in both plasmids.